De-ID® Update February 2003

Project Team

- Development by the Center for Biomedical Informatics
  - Paul Hanbury
  - Melissa Saul
  - Greg Cooper, M.D., Ph.D.
  - Bruce Buchanan, Ph.D.
  - Mehmet Kayaalp M.D., M.S.
  - Wendy Chapman, Ph.D.

- Implementation by the Clinical Research Informatics Service (CRIS)

De-ID® Mechanics

- De-ID uses a set of heuristics to identify the presence of any of the 18 HIPAA identifiers within the text.
- Supplemental dictionaries of geographic locations, hospital names, and popular names found in the U.S. Census are also used to locate identifiable text.
- The UMLS Methatheasurus is utilized to ensure that words or phrases that may be medical terms containing proper names are preserved.
- De-ID replaces the identifiable text with specific tags. Names found multiple times in the report are consistently replaced with the same tag to improve readability of the report .

De-ID® Software Development

- Version 1.0 – basic scrubbing of names
- Version 2.0 (May 2001) – handling of for HIPAA defined identifiers as related to text for six commonly used MARS documents
- Version 3.0 (October 2001) – improvements based on CBMI audit
- Version 3.3 (January 2002) – additions to accommodate non-sentence structured documents such as pathology
- Version 4.0 (August 2002) – improvements based on 9 month use; assignment of unique patient ID
- Version 4.14 (January 2003) – refinements in handling of common words and phrases

The Use of De-ID®

Clinical Research within the Schools of the Health Sciences

- Epidemiology, GSPH
- Health Info Mgmt, SHRS
- Biomedical Informatics, SOM
- Critical Care Medicine, SOM
- ENT, SOM
- Medicine, SOM
- Neurology, SOM
- Orthopedic Surgery, SOM
- Pathology, SOM
- Radiology, SOM
- Outcomes Center, School of Pharmacy
- Acute and Tertiary Care, School of Nursing
- Center for Chronic Diseases, School of Nursing
- Shared Pathology Informatics Network (SPIN) – NCI Funded Collaboration

The Use of De-ID

UPMC Projects

- UPMC/RAND Women's Health Initiative
- Electronic Health Record Project

Other Opportunities

- Boston area biotechnology company involved in tissue-typing analysis
- VA Healthcare system coordinating a multi-center study
- Healthcare vendor utilizing existing data for benchmarking purposes

De-ID Usage Statistics

	FY02	FY03YTD
• Datasets de-identified	124	85
• IRB Exempt Projects	51	45
• Formal Audits	3	1

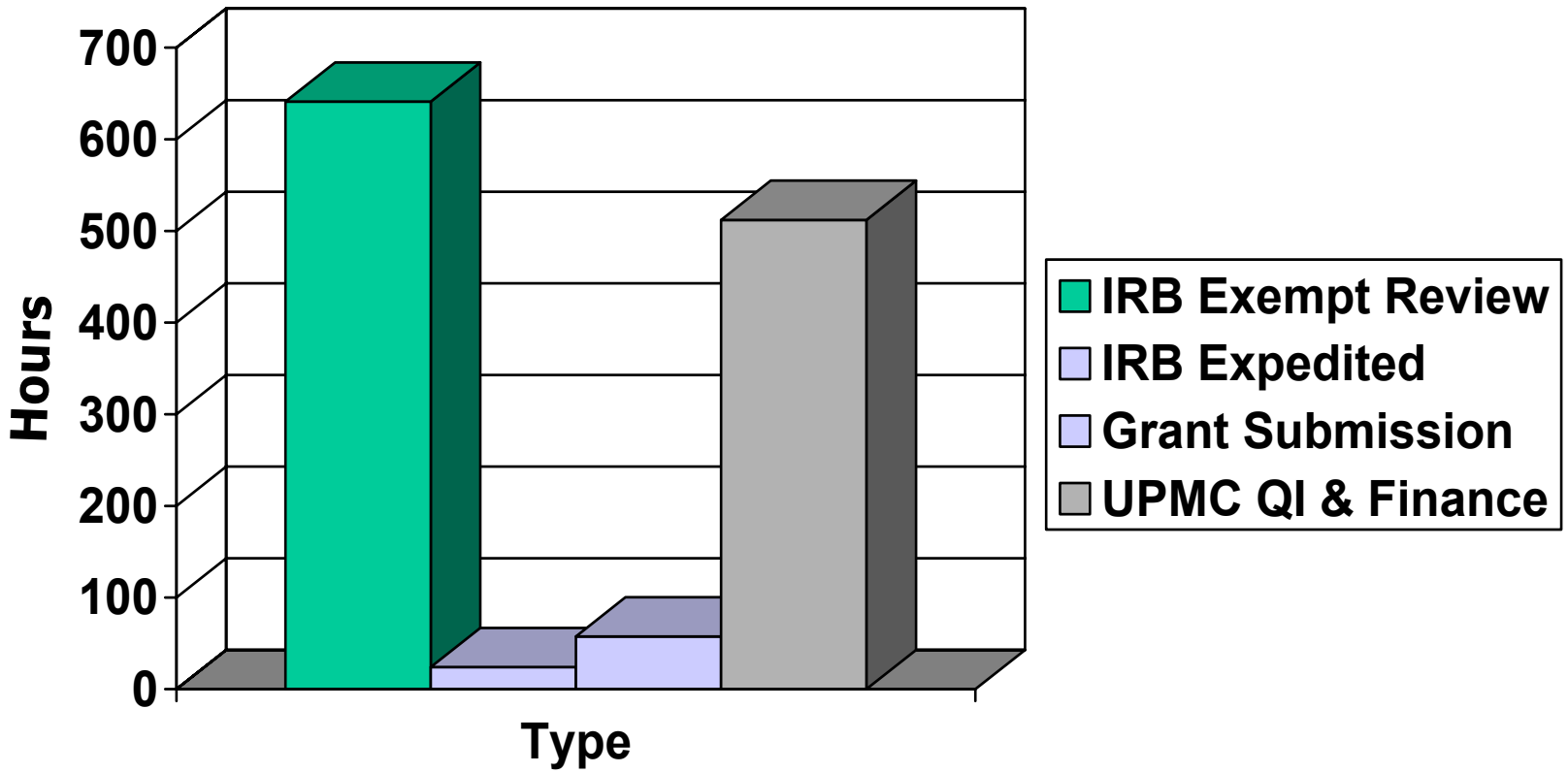
De-ID in IRB Exempt Studies

IRB Application Form, Page 1:

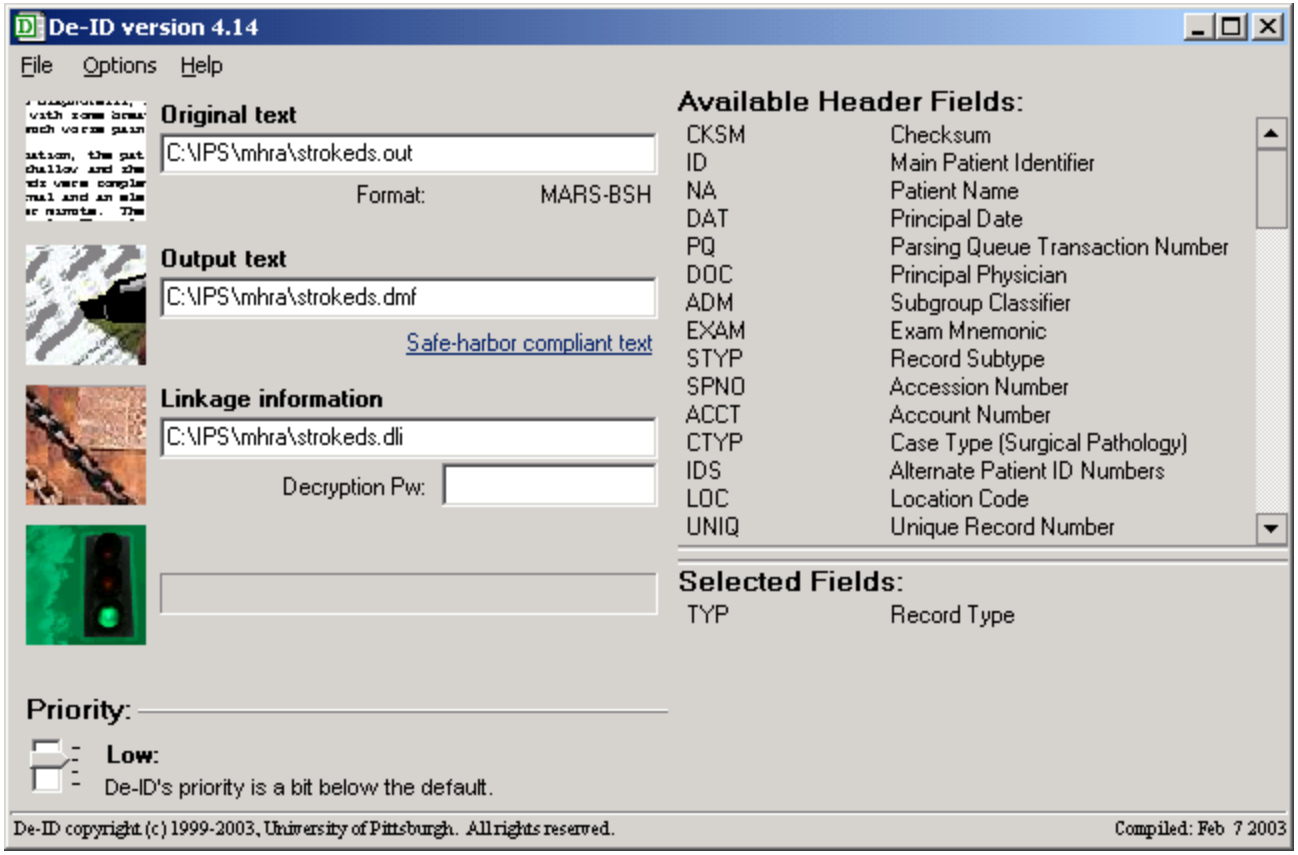
How will the information be recorded?

*Information electronically de-identified by a computerized system over which research team has no access (e.g., **UPMC (sic) De-ID** program; other program (name):*

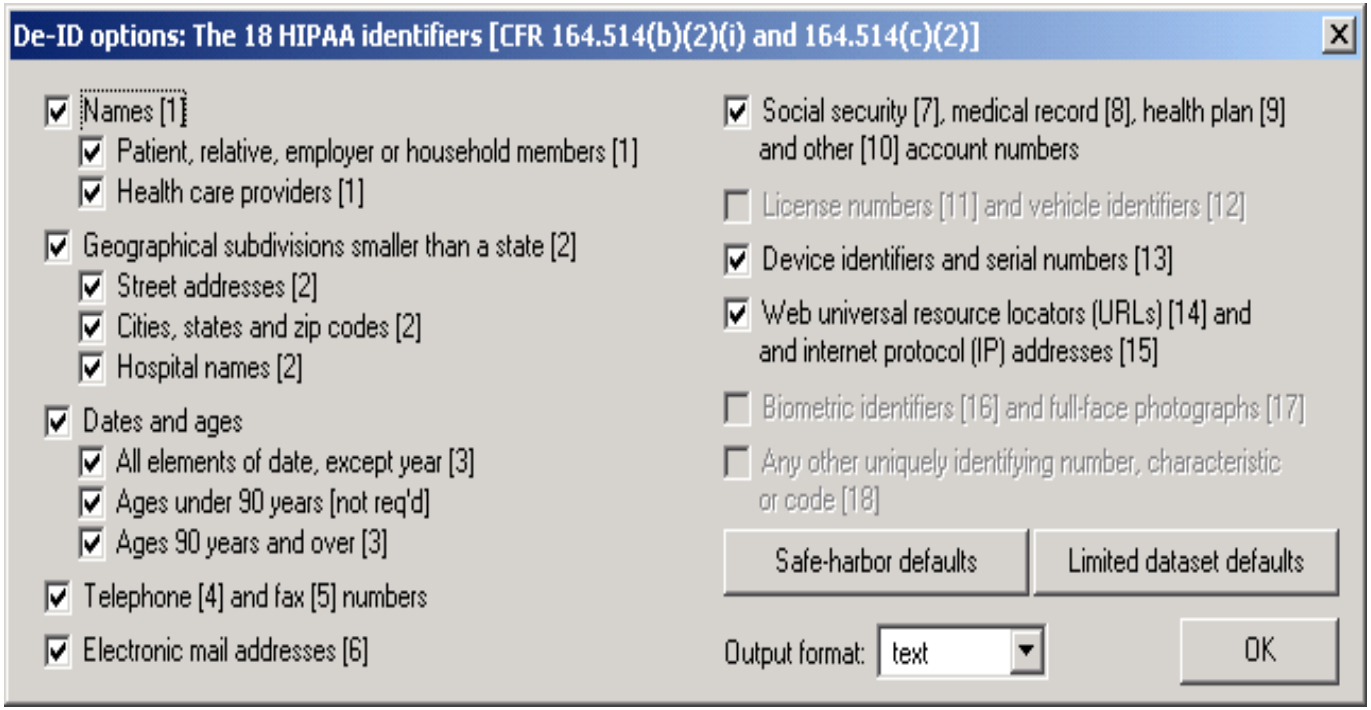
CRIS Projects by Type – FY02



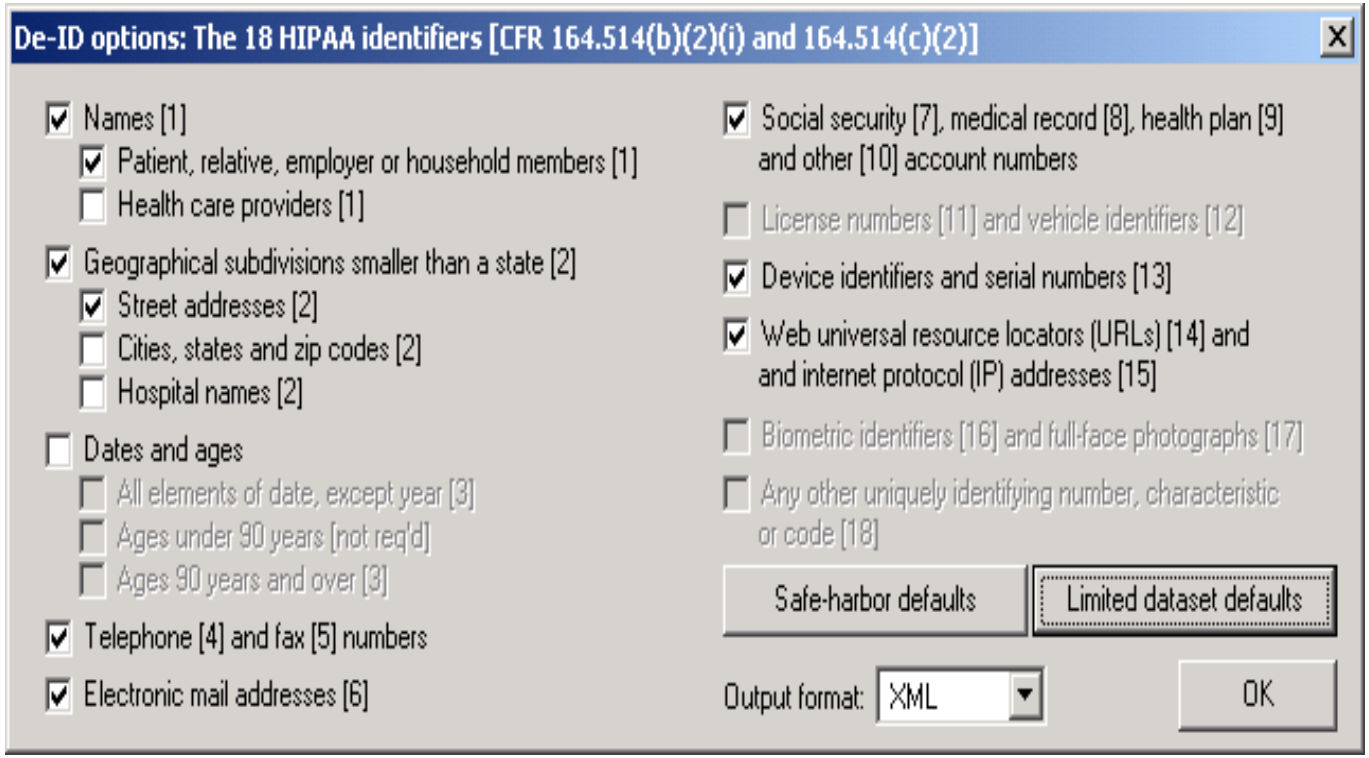
De-ID® Application



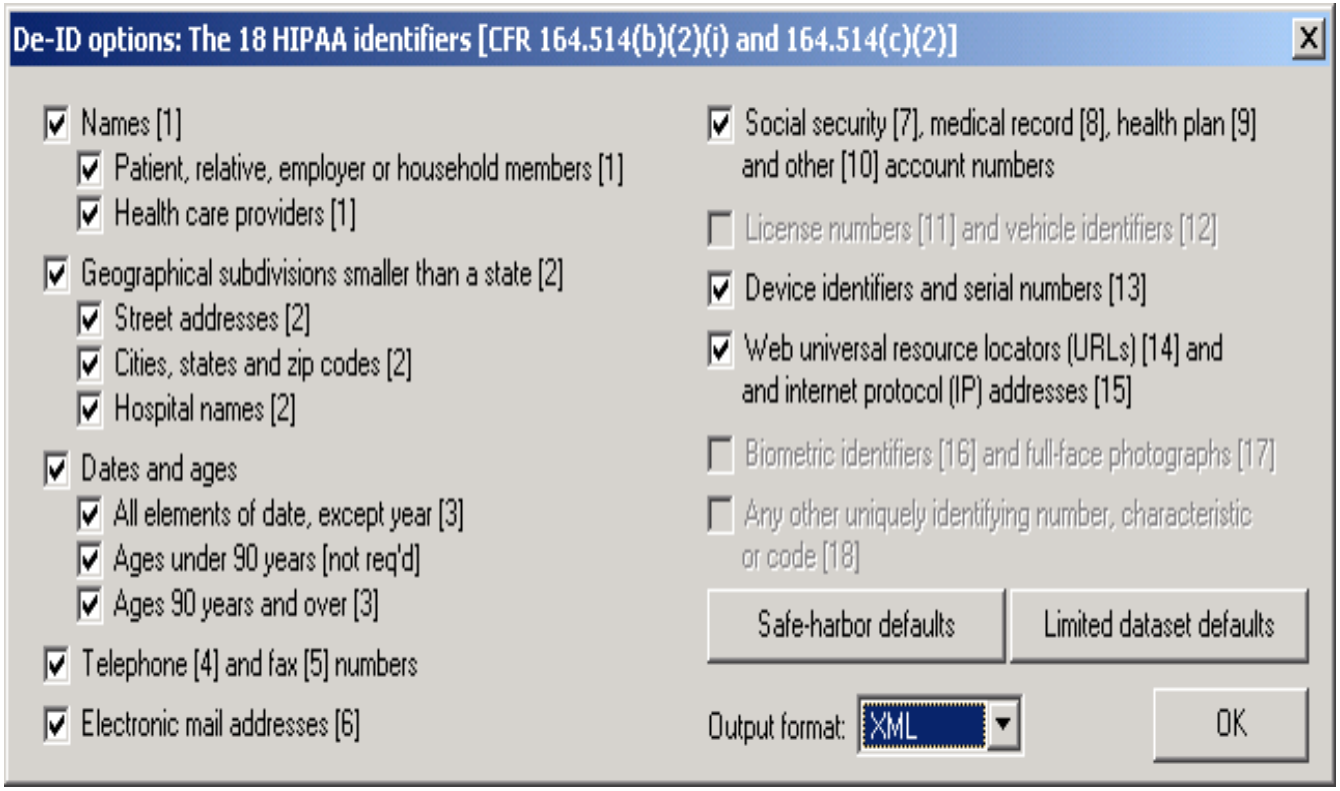
De-ID Options



DE-ID Limited Dataset Defaults



XML Formatted Output Option



Sample De-ID Report

REASON FOR ADMISSION:

This is a **AGE[in 60s]-year-old male with a past medical history significant for hypertension, who developed a throat discomfort while swimming. The patient was admitted on **DATE[May 17 2001]. The patient was admitted with crescendo angina and positive troponin 1.

HOSPITAL COURSE:

The patient was admitted to a monitored bed, cardiology was consulted and the patient was seen. The patient was placed on aspirin and heparin. The patient was also placed on a beta-blocker. Dr. **NAME[VVV UUU] from cardiology saw the patient and recommended a cardiac catheterization. On **DATE[May 18 2001], the patient had a cardiac catheterization,

Formal Audit Studies

- Initial evaluation study in 2001 of 350 reports from UPMC PUH
- Study of 967 surgical pathology reports from nine UPMC-HS hospitals from 2000
- Study of 1000 surgical pathology reports from UPMC PUH from sign outs of years 1985, 1990, 1995, 2000
- Current study of 750 reports from UPMC PUH evaluated by experienced physicians

Refinement Process

- Goal: to identify and remove all phone numbers
  - Formats
    - 3-3-4
    - 3-4
    - 3*7
    - 7 with a .
    - 10 with a .
- Issues
  - ✓800-1000 cc of fluid
  - ✓FAX # *4125551212